CIRJE-F-1124

Mechanism Design with General Ex-Ante Investments

Hitoshi Matsushima
The University of Tokyo

Shunya Noda
University of British Columbia

April 2020

Mechanism Design with General Ex-Ante Investments[1]

Hitoshi Matsushima[2]*
University of Tokyo

Shunya Noda[3]
University of British Columbia

First Draft: May 30th, 2016

This Version: September 5th, 2019

Abstract

We investigate the general mechanism design problems in which agents take ex-ante hidden actions (or investments) that influence state distribution. We show that the variety of action choices drastically shrinks the set of mechanisms that induce targeted actions in the sense that there is a one-to-one tradeoff between the dimensionality of action space and that of payment rules with which the targeted action profile is taken in an equilibrium. This result comprehends the observations made by previous works. When agents can take unilateral deviations to change the state distribution in various directions, we have equivalence properties with respect to ex-post payoffs, payments, and revenues. In particular, the *pure-VCG mechanism*, the simplest form of the canonical VCG mechanism, becomes the only mechanism that induces an efficient action profile. Contrarily, the popular pivot mechanism generically fails to induce efficient actions, even when the action space is one-dimensional.

Keywords: hidden action, hidden information, VCG-necessity, equivalence

JEL Classification: C72, D44, D82, D86

[1] This paper supersedes "Mechanism Design in Hidden Action and Hidden Information: Richness and Pure-VCG."

[2]* Corresponding author. Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; hitoshi[at]e.u-tokyo.ac.jp.

[3] Vancouver School of Economics, University of British Columbia, 6000 Iona Dr, Vancouver, BC, V6T 1L4, Canada; shunya.noda[at]gmail.com.

1. Introduction

The larger the action space, the more serious the hidden action problem. Large action spaces restrict the set of mechanisms/contracts that successfully incentivize agents to take targeted actions. Such restrictions on mechanisms may make it impossible to (i) solve other problems the mechanism designer faces (e.g., hidden information, participation constraints) or (ii) satisfy the properties demanded for the mechanism (e.g., efficiency, non-negative revenue). To identify a mechanism that satisfies all of the required properties at the same time, we need to characterize the relationship between the largeness of an action space and the smallness of a set of mechanisms that induce targeted actions.

This paper studies a general class of mechanism design problem in which a hidden action problem and hidden information problem coexist. Agents initially take general ex-ante actions/ investments, which determines the probability distribution over the state of the world. Once the state is realized, the mechanism (which the central planner commits beforehand) determines the allocation and payment in a state-contingent manner. Neither actions nor states (agents' private types) are assumed to be observable. Hence, the central planner must carefully design a mechanism to (i) *induce* the targeted action profile to generate a targeted state distribution (i.e., the targeted action profile comprises a Nash equilibrium of the game implied by the mechanism), and (ii) incentivize the truthful reporting of agents' private types so as to implement the targeted allocation rule.

Many mechanism design problems, such as principal–agent problems, partnerships, auctions, and public good provision, have such a structure. Practically, agents in such problems often have opportunities to take ex ante actions, such as information acquisition, R&D investment, patent control, standardization, M&A, rent-seeking, positive/negative campaigns, product differentiation, entry/exit decisions, preparation of infrastructure, and headhunting. Each of these actions influences the state of the world in a different way. Furthermore, participants might be able to take multiple kinds of actions. To study these problems comprehensively, we consider a very general class of action spaces.

Our focus is on the relationship between the largeness of the action space and the smallness of the set of mechanisms that induce a targeted action profile. The largeness of an action space can be measured by the dimensionality of the set of state distributions generated from unilateral deviations. The smallness of a set of mechanisms can be measured by the dimensionality of payment rules with which the targeted action profile is induced (given a targeted allocation rule). We characterize the relationship between these two, which we then use to comprehend a variety of issues related to the hidden action and information problems.

The main characterization theorem in this study articulates the one-to-one tradeoff between the dimensionality of action space and that of the set of well-behaved mechanisms. We show that, as the dimensionality of action space increments by one, the upper-bound of the dimensionality of the set of payment rules inducing the targeted action profile decrements by one (Theorem 1). This bound is tight with additional assumptions, with which the global optimality of action choices is implied by the local optimality (Theorem 2). Theorems 1 and 2 indicate that, even when a mechanism induces a targeted action profile with an action space, once there arises an additional, different kind of actions, the mechanism *generically* fails to induce the same actions.

As Theorems 1 and 2 imply, the presence of hidden actions imposes a severe restriction on the design of desirable mechanisms. As corollaries, we complement a number of previous results related to hidden action and information problems.

First, we show that the popular *pivot mechanism* (due to Green and Laffont, 1979), which aligns an agent's payoff with his marginal contribution to social welfare, generically fails to induce efficient actions (Proposition 2). Intuitively, this is because each agent is tempted to weaken other agents to exaggerate his marginal contribution. This generalizes Krähmer and Strausz's (2007) observation, which is made for the case of one-dimensional investment space.

We also derive a uniqueness theorem for the case of large action spaces. As the dimensionality of the action space increases, the dimensionality of the set of desirable payment rules decreases. In an extreme case, each agent can unilaterally change the state distribution in full-dimensional directions. We say the action space is *rich* in such a case. Krähmer and Strausz (2007) and Athey and Segal (2013) prove that, unlike the pivot mechanism, *the pure-VCG mechanism* always induces an efficient action profile. The pure-VCG mechanism is the simplest form of canonical Vickrey–Clarke–Groves (VCG) mechanisms,[4] which gives each agent the welfare of the other agents and then imposes a fixed monetary fee. We newly show the necessity of pure-VCG in that any mechanism other than pure-VCG fails to induce an efficient action profile if the action space

[4] Vickrey (1961), Clarke (1971), and Groves (1973).

is sufficiently large. Accordingly, when the action space is rich, pure-VCG is the unique mechanism inducing an efficient action profile (Theorem 4).

We further obtain a full characterization theorem for general targeted actions. If an action space is rich, the set of mechanisms that induce an efficient action profile becomes one-dimensional: such mechanisms are unique up to constants (Theorem 3). Therefore, if the planner desires to induce a targeted action, there will be no degree of freedom for designing a payment rule to incentivize truthful reporting. This is a new notion of the "revenue-equivalence principle," which extends Theorem 3 of Holmström and Milgrom (1987) (proven for a principal–agent problem) to a general mechanism design problem.

Finally, we also study an environment that have *independent types*: the influence of each agent's action is restricted to his own private type distribution. Even in such a case, the variety of actions restricts the set of mechanisms that solve the hidden action problem. Parallel to Theorem 3, assuming a large action space, we derive a characterization theorem—all payment rules that induce a targeted action profile are interim equivalent, rather than ex-post (Theorem 5). In particular, a mechanism induces an efficient action profile if and only if it is the *expectation-VCG mechanism*: each agent's expected payment is equal to that from a VCG mechanism (Theorem 6).[5] Our theorems complement the observations made by Rogerson (1992), Bergemann and Välimäki (2002) and Hatfield, Kojima, and Kominers (2015). For further detail, we direct the reader to Section 7.

The rest of the manuscript is structured as follows. Section 2 describes the model. Section 3 reviews the literature. Section 4 studies the hidden action problem. Section 5 analyzes the hidden information problem. Section 6 studies an efficient mechanism design. Section 7 considers the case without externality and compares the result with those of previous works. Section 8 concludes the paper.

[5] Matthews (1984), Hausch and Li (1991), and Tan (1992) show that the first- and second-price auctions provide each agent with the same incentive in hidden action in symmetric environments, in which the first-price auction is an indirect implementation of an expectation-VCG mechanism and the second-price auction is equivalent to the pivot mechanism. Tan (1992), Stegeman (1996), and Arozamena and Cantillon (2004) also show that, with private values, the second-price auction (the pivot mechanism) induces an efficient hidden action. This study provides a comprehensive understanding of all such previous works.

2. Literature Review

Some papers have demonstrated the desirability of simple contracts in complex environments. Holmström and Milgrom (1987), for instance, study the moral hazard problem with a principal–agent model. They show that when an agent's action space is sufficiently large, there is a unique contract inducing the targeted effort level. Carroll (2014) proves that when the principal faces ambiguity about the range of activities, the optimal robust contract must be linear with respect to the resultant output. This paper studies a general class of mechanism design problem, of which the principal–agent model is a special case. For such an environment, we show that pure-VCG, which is the simplest form of the VCG mechanism, is a unique fully efficient mechanism when the action space is sufficiently large (i.e., it satisfies the richness condition).

The literature has also shown that, when the scope of action spaces is sufficiently limited, we can design efficient incentive mechanisms by tailoring the payment rule to detailed specifications. We can even achieve either full surplus extraction or budget balance (e.g., Matsushima, 1989; Legros and Matsushima, 1991; Williams and Radner, 1995; Obara, 2008). However, when agents can take various actions, a mechanism's dependence on the detail tempts each agent to deviate. We study not only the case with extremely large or extremely small action spaces but also intermediate ones. Our Theorem 1 implies that even if agents can only take a few kinds of actions, the set of mechanisms inducing targeted actions shrinks drastically.

It is well-known that when the state space is "large," the set of incentive compatible mechanism becomes "small." Green and Laffont (1977, 1979) and Holmström (1979) show that, in hidden information environments with differentiable valuation functions, differentiable path-connectedness, and private values, VCG mechanisms are the only efficient mechanisms. This study reconsiders VCG mechanisms from the viewpoint of the hidden action problems and shows that only pure-VCG mechanisms, which are special cases of VCG, can induce an efficient action profile.[6] In this sense, our characterization theorem is a new version of the VCG-necessity

[6] Hausch and Li (1993) and Persico (2000) are relevant to this point. They demonstrate that first- and second-price auctions provide different incentives for information acquisition that make the other agents' valuation more accurate. Taking the mechanism design approach rather than comparing special auction formats, we explain that the difference in the induced action profile originates from the difference in ex-post payoffs between these auction formats.

theorem. Importantly, our equivalence theorems rely on assumptions that are completely different from Green and Laffont (1977, 1979) and Holmström (1979).

3. Model

Consider a setting with one central planner and n agents indexed by $i \in N = \{1,2,\ldots,n\}$. The central planner commits to a mechanism at the beginning. Observing the committed mechanism, agents choose ex-ante actions that determine the state distribution. After observing the realized private types, agents participate in the mechanism and report their private types. To be more precise, the "game" proceeds as follows.

Stage 1: The central planner commits to a *mechanism* defined as (g, x), where Ω denotes the set of states, A denotes the set of allocations, $g: \Omega \to A$, and $x \equiv (x_i)_{i\in N} : \Omega \to R^n$. We assume that Ω and A are finite. We call g and x the *allocation rule* and *payment rule*, respectively.[7]

For each $i \in N$, we call a pair of an allocation rule and a payment rule for agent i (i.e., (g, x_i)) a *mechanism for agent* i. Slightly abusing notation, x_i is sometimes regarded as a $|\Omega|$-dimensional vector (i.e., $x_i = (x_i(1),\ldots,x_i(|\Omega|)) \in R^{|\Omega|}$). We also denote $x = (x_1,\ldots,x_n) = (x_i(\omega))_{i\in N,\omega\in\{1,\ldots,|\Omega|\}} \in R^{n|\Omega|}$.

Stage 2: Each agent $i \in N$ selects a hidden action $b_i \in B_i$, where B_i denotes the set of all actions for agent i. The cost function of agent $i's$ action choice is given by $c_i : B_i \to R_+$. We assume that there is a no-effort option $b_i^0 \in B_i$ such that $c_i(b_i^0) = 0$. Let $B \equiv \times_{i\in N} B_i$ and $b \equiv (b_1,\ldots,b_n) \in B$.

[7] This study assumes that the central planner commits to the mechanism *before* agents take actions. Without the assumption of commitment, the desired outcomes, such as efficiency, may be unachievable due to time inconsistency. Consider a central planner who commits to a mechanism *after* the agents' action choices. In this case, the central planner prefers the pivot mechanism because it yields greater revenue than the mechanism proposed in this study (the pure-VCG mechanism). However, the pivot mechanism typically fails to induce any efficient action profile (see Proposition 2). Anticipating that the use of a pivot mechanism, agents are willing to select inefficient actions.

Stage 3: The state $\omega = (\omega_0, \omega_1, \cdots, \omega_n)$ is randomly drawn from the probability function $f(\cdot|b) \in \Delta(\Omega)$, where $b \in B$ is the action profile selected at stage 2. We call ω_0 a *public signal* and ω_i a *type* for each agent $i \in N$. Let Ω_0 denote the set of all public signals and Ω_i denote the set of all types of agent i. Let $\Omega \equiv \times_{i \in N \cup \{0\}} \Omega_i$, and $\Omega_{-i} \equiv \times_{N \cup \{0\} \setminus \{i\}}$. Agent i observes her type ω_i and public signal ω_0 but cannot observe ω_{-i} at this stage.

Stage 4: Each agent $i \in N$ reports $\widetilde{\omega}_i \in \Omega_i$ about her type. Afterward, all agents, as well as the central planner, observe the public signal ω_0. According to the profile of the agents' reports $\widetilde{\omega}_{-0} = (\widetilde{\omega}_i)_{i \in N}$ and the observed public signal ω_0, the central planner determines the allocation $g(\omega_0, \widetilde{\omega}_{-0}) \in A$ and the side payment vector $x(\omega_0, \widetilde{\omega}_{-0}) \in R^n$. The resultant payoff of agent i is given by

$$v_i(g(\omega_0, \widetilde{\omega}_{-0}), \omega) - x_i(\omega_0, \widetilde{\omega}_{-0}) - c_i(b_i).$$

Here, we assume that each agent's payoff function is quasi-linear and risk-neutral, and the cost of the agent's action choice is additively separable.

Figure 1 describes the timeline of the model. We will study the way to implement the targeted action profile and allocation rule, (b, g).

Remark 1: Without loss of generality, we can focus on the revelation mechanisms in which each agent reports only her type, because the central planner attempts to induce a pure action profile (see Proposition 1 in Obara, 2008 for the revelation principle with ex-ante actions).[8]

[8] Obara (2008) also argues that, if the central planner attempts to induce a mixed action profile, we need to consider mechanisms in which each agent reports not only her type but also her selection of pure action. In a model with finite action spaces, by carefully mixing agents' actions, Obara (2008) generates a correlation between (b_i, ω_i) and (b_{-i}, ω_{-i}) and then applies the technique of Crémer and McLean (1985, 1988) to the incentives in revelations. However, the argument by Obara needs to make extremely large side payments and utilize very detailed knowledge about specifications.

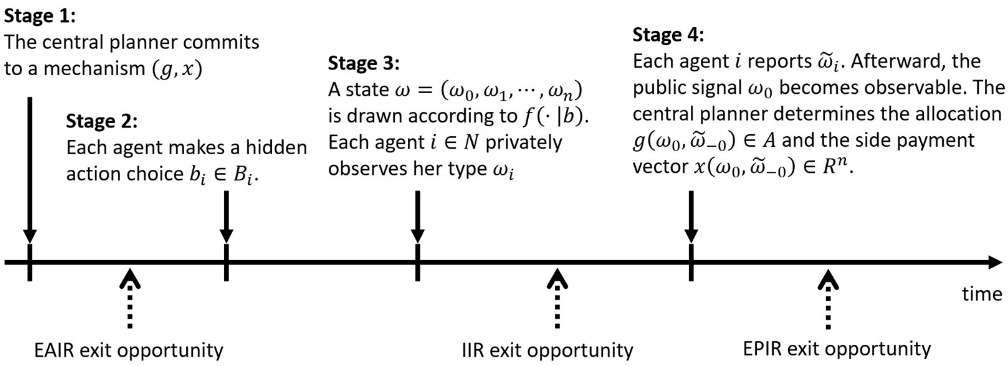

Figure 1: Timeline with Hidden Action and Hidden Information

4. Hidden Action

This section focuses on incentives in the hidden action problem. Specifically, we assume that agents truthfully report the state $\omega = (\omega_0, \omega_1, \cdots, \omega_n)$ in stage 4, and we characterize the mechanism that induces a targeted action profile. All of the results described in this section would be true even if the state were publicly observable.

Definition 1 (Inducibility): A mechanism for agent i, (g, x_i), is said to *induce* an action profile $b \in B$ *for agent* i if b_i is a best response to b_{-i}, i.e., for all $b_i' \in B_i$,

$$E[v_i(g(\omega), \omega) - x_i(\omega)|b] - c_i(b_i) \geq E[v_i(g(\omega), \omega) - x_i(\omega)|b_i', b_{-i}] - c_i(b_i'),$$

where $E[\cdot\,|b]$ denotes the expectation operator given that action profile b is taken: for every function $\xi : \Omega \to R$, $E[\xi(\omega)|b]$ is defined as follows:

$$E[\xi(\omega)|b] \equiv \sum_{\omega \in \Omega} \xi(\omega) f(\omega|b).$$

A mechanism (g, x) *induces* an action profile b if (g, x_i) induces b for every $i \in N$, i.e., b is a Nash equilibrium of the game implied by the mechanism (g, x).

Define $X_i(b, g)$ as

$$X_i(b,g) \equiv \{\tilde{x}_i \in R^{|\Omega|}: (g,\tilde{x}_i) \text{ induces } b \text{ for } i\}.$$

In words, $X_i(b,g)$ denotes the set of all payment rules for agent i that induce targeted action profile b along with allocation rule g.

$X_i(b,g) \neq \emptyset$ if and only if there exists a function $u_i : \Omega \to R$ such that

$$b_i \in \underset{b_i' \in B_i}{\operatorname{argmax}}\{E[u_i(\omega)|b_i', b_{-i}] - c_i(b_i')\}$$

for all $i \in N$. In such a case, $\tilde{x}_i(\omega) = v_i(g(\omega),\omega) - u_i(\omega)$ satisfies $\tilde{x}_i \in X_i(b,g)$. Accordingly, if the planner has an objective function and b is an optimal solution, then $X_i(b,g)$ is non-empty.[9] For example, if (b,g) maximizes the social welfare (i.e., (b,g) is fully efficient; see Section 6), then $X_i(b,g)$ is always non-empty. Hereinafter we focus on (b,g) such that $X_i(b,g)$ is non-empty for every $i \in N$.

The *dimension* of a vector space $L \subset R^h$, denoted by $\dim L$, is defined as a cardinality of a basis of L. The *dimension* of a general set $L' \subset R^h$, denoted by $\dim L'$, is defined as the minimal dimension of vector spaces containing L'. We regard $\dim X_i(b,g)$ as a measure of the largeness of $X_i(b,g)$; i.e., the variety of payment rules x such that the associated mechanisms (g,x) induce the targeted action profile b.

To measure the availability of unilateral deviations, we introduce the following concept.

Definition 2 (Deviation path): A mapping $\beta_i : [-1,1] \to B_i$ is said to be a *deviation path of* $(i,b) \in N \times B$ if

$$\beta_i(0) = b_i,$$
$$[\beta_i(\alpha) = \beta_i(\alpha')] \Rightarrow [\alpha = \alpha'],$$
$$c_i(\beta_i(\alpha)) \text{ is differentiable in } \alpha \text{ at } \alpha = 0,$$

and there exists a *tangent* of β_i at b, which is denoted by $t(\beta_i) \in R^{|\Omega|}$, where

[9] For example, $X_i(b,g)$ is empty if b_i is "dominated": there exists $b_i' \in B_i$ such that $f(\cdot\,|b) = f(\cdot\,|b_i', b_{-i})$ and $c_i(b_i') < c_i(b_i)$. In such a case, as long as the other agents take b_{-i}, for any payment rule x_i, agent i does not take b_i; thus, it is impossible to induce b. However, if we assume that the planner maximizes her objective function (e.g., revenue or efficiency) through mechanism design, she has no reason to induce such actions.

$$t(\beta_i) \equiv \lim_{\alpha \to 0} \frac{f(\cdot\,|\beta_i(\alpha), b_{-i}) - f(\cdot\,|b)}{\alpha}.$$

A deviation path is an ordered subset of the action space. Since each point on the deviation path is just an action, an agent can always make a local unilateral deviation "along with" the deviation path; i.e., to take $\beta_i(\alpha) \in B_i$ with $\alpha \neq 0$. A mechanism (g, x) induces b only if it does not incentivize such a deviation.

Each deviation path can be interpreted as a class of actions. For example, if an agent can make two different classes of ex-ante actions, say, R&D investment and preemptive behavior, he can make a deviation in R&D investment and preemptive behavior, whichever they want. To induce a targeted action profile, the mechanism must discourage agents from taking a deviation in both of these two actions, simultaneously. In this sense, the set of mechanisms that successfully induce a targeted action in this case is an intersection of the set of mechanisms that prevent agents from taking a deviation (i) in R&D investment, and (ii) preemptive behavior. By considering deviations along with each deviation path, we analyze (i) and (ii), separately.

Figure 2a illustrates an example of deviation path β_i. We assume $|\Omega| = 3$, and the triangle represents the probability simplex $\Delta(\Omega)$. The blue point represents $f(\cdot\,|b)$. Each point on the orange curve corresponds to a probability distribution that can be generated by agent i's unilateral deviation along deviation path β_i. The tangent of β_i at b, $t(\beta_i)$, is depicted as the green-dotted arrow.

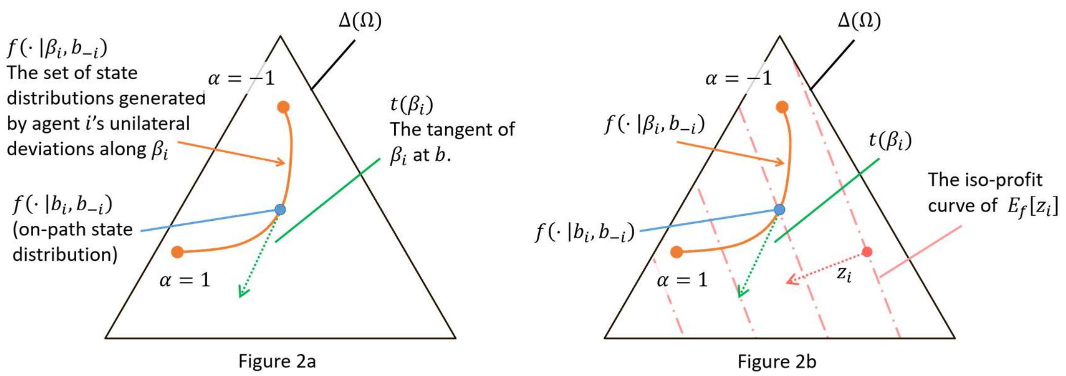

Figure 2: Deviation path. If z_i is not perpendicular to $t(\beta_i)$, agent i has an incentive to make a unilateral deviation along with β_i (i.e., increase or decrease α slightly).

Fix an arbitrary payment rule for agent i, $x_i \in X_i(b,g)$. We define

$$Z_i(b,g,x_i) \equiv \{z_i \in R^{|\Omega|} : x_i + z_i \in X_i(b,g)\}.$$

Clearly, we have

$$\dim X_i(b,g) = \dim Z_i(b,g,x_i).$$

Proposition 1 shows that each deviation path constrains $Z_i(b,g,x_i)$ (and $X_i(b,g)$) by imposing one equality constraint.

Proposition 1: *Suppose that* (g,x_i) *induces* b *for agent* i *and there exists a deviation path* β_i *of* $(i,b) \in N \times B$. *Then, we have* $z_i \in Z_i(b,g,x_i)$ *only if*

(1) $$z_i \cdot t(\beta_i) = 0.$$

Proof: Consider an arbitrary $z_i \in Z_i(b,g,x_i)$. Since (g,x_i) induces b for agent i, the following first-order condition is necessary for preventing agent i from locally deviating along β_i:

(2) $$\frac{d}{d\alpha}\{E[v_i(g(\omega),\omega) - x_i(\omega)|\beta_i(\alpha),b_{-i}] - c_i(\beta_i(\alpha))\}\Big|_{\alpha=0} = 0.$$

Similarly, $(g, x_i + z_i)$ must satisfy the following first-order condition:

(3) $$\frac{d}{d\alpha}\{E[v_i(g(\omega),\omega) - x_i(\omega) - z_i(\omega)|\beta_i(\alpha),b_{-i}] - c_i(\beta_i(\alpha))\}\Big|_{\alpha=0} = 0.$$

Subtracting (3) from (2), we obtain

$$\frac{d}{d\alpha}E[z_i(\omega)|\beta_i(\alpha),b_{-i}]\Big|_{\alpha=0} = 0,$$

or equivalently, $z_i \cdot t(\beta_i) = 0$.

Q.E.D.

Proposition 1 implies that a payment rule $\tilde{x}_i$ for agent i is never included in $X_i(b,g)$ if the difference $z_i \equiv \tilde{x}_i - x_i$ is not perpendicular to the tangent of the deviation path β_i (i.e., $z_i \cdot t(\beta_i) \neq 0$). Since (g, x_i) induces b, agent i is (approximately) indifferent for local deviations along the deviation path β_i in (g, x_i). If $(g, x_i + z_i)$ induces b, agent i must also be indifferent for local deviations along β_i in $(g, x_i + z_i)$. This implies that the expected value of z_i is (approximately) unchanged even if agent i locally deviates. Figure 2b shows a case of $z_i \cdot t(\beta_i) > 0$, where agent i has an incentive to take α slightly larger than zero. Accordingly, $(g, x_i + z_i)$ fails to induce b.

In general, there exist multiple deviation paths whose tangents are linearly independent. In such a case, the equality condition (1) must be satisfied for all deviation paths. Accordingly, one deviation path decreases the upper-bound of the dimensionality of $X_i(b,g)$ by one as follows:

Theorem 1: *Suppose there exist K_i deviation paths of (i,b), denoted by β_i^1, ..., and $\beta_i^{K_i}$, such that the respective tangents $t(\beta_i^1)$, ..., and $t\left(\beta_i^{K_i}\right)$ are linearly independent. Then, we have*

$$\dim X_i(b,g) \leq |\Omega| - K_i.$$

Proof: Proposition 1 implies

$$Z_i(b,g,x_i) \subset \left\{z_i \in R^{|\Omega|} : z_i \cdot t\left(\beta_i^k\right) = 0 \text{ for all } k = 1,2,\cdots,K_i\right\}.$$

Hence,

$$\dim Z_i(b,g,x_i) \leq \dim\left\{z_i \in R^{|\Omega|} : z_i \cdot t\left(\beta_i^k\right) = 0 \text{ for all } k = 1,2,\cdots,K_i\right\}.$$

Since $t(\beta_i^1)$, …, and $t(\beta_i^{K_i})$ are linearly independent,

$$\dim\left\{z_i \in R^{|\Omega|} : z_i \cdot t\left(\beta_i^k\right) = 0 \text{ for all } k = 1,2,\cdots,K_i\right\} \leq |\Omega| - K_i.$$

Therefore,

$$\dim Z_i\,(b,g,x_i) \leq |\Omega| - K_i.$$

Hence, we have $\dim X_i(b,g) = \dim Z_i(b,g,x_i) \leq |\Omega| - K_i$ as desired.

Q.E.D.

Theorem 1 only gives an upper-bound on the dimensionality of $X_i(b,g)$. This is because Proposition 1 and Theorem 1 only consider local deviations along with deviation paths. Agents may also make a deviation to a distant action, and this opportunity further constrains the set of payment rules that induce b. Consequently, $\dim X_i(b,g)$ could be smaller than $|\Omega| - K_i$.

When local optimality of action choices implies global optimality, the number of deviation paths provides a tight characterization of the set of inducible mechanisms. Theorem 2 exhibits this result.

Theorem 2: *Suppose that for every $b_i' \in B_i$, there exists a deviation path β_i that satisfies the following two conditions:*

(4) $$t(\beta_i) = f(\cdot\,|b_i', b_{-i}) - f(\cdot\,|b_i, b_{-i}),$$

(5) $$\left.\frac{d}{d\alpha} c_i\big(\beta_i(\alpha)\big)\right|_{\alpha=0} \leq c_i(b_i') - c_i(b_i).$$

Suppose also that K_i is the maximum number of deviation paths of (i,b), denoted by β_i^1, ..., and $\beta_i^{K_i}$, such that the respective tangents $t(\beta_i^1)$, ..., and $t\big(\beta_i^{K_i}\big)$ are linearly independent. Then,

$$\dim X_i(b,g) = |\Omega| - K_i.$$

Proof: Suppose that b_i' is a profitable deviation from b_i given a mechanism (g,x); i.e.,

(6) $$E[v_i(g(\omega),\omega) - x_i(\omega)|b_i', b_{-i}] - c_i(b_i') > E[v_i(g(\omega),\omega) - x_i(\omega)|b] - c_i(b_i).$$

By assumption, there exists a deviation path β_i that satisfies (4) and (5). By (5) and (6), we have

$$E[v_i(g(\omega),\omega) - x_i(\omega)|b_i', b_{-i}] - E[v_i(g(\omega),\omega) - x_i(\omega)|b] > \left.\frac{d}{d\alpha} c_i\big(\beta_i(\alpha)\big)\right|_{\alpha=0}.$$

By (4), we have

(7) $$\frac{d}{d\alpha}E[v_i(g(\omega),\omega)-x_i(\omega)|\beta_i(\alpha),b_{-i}]\Big|_{\alpha=0}$$
$$=E[v_i(g(\omega),\omega)-x_i(\omega)|b_i',b_{-i}]-E[v_i(g(\omega),\omega)-x_i(\omega)|b].$$

By (6) and (7), we have

$$\frac{d}{d\alpha}\{E[v_i(g(\omega),\omega)-x_i(\omega)|\beta_i(\alpha),b_{-i}]-c_i(\beta_i(\alpha))\}\Big|_{\alpha=0}>0.$$

Accordingly, when there exists a profitable deviation, there exists a deviation path such that a local deviation along with the path is also profitable.

Fix a mechanism (g, x_i) that induces the action profile b for i. By the above argument, a mechanism $(g, x_i + z_i)$ induces b for i if and only if no local deviation along any deviation paths is profitable; i.e., z_i is perpendicular to all tangents of the deviation paths. Since there are exactly K_i deviation paths whose tangents are linearly independent, the set of such z_i's is exactly $|\Omega| - K_i$ dimensional.

Q.E.D.

A sufficient condition for (4) and (5) is the availability of "intermediate" actions in the following sense. Suppose that two actions b_i and b_i' are available. Then, for all $\lambda \in [0,1]$, there exists another action $\tilde{b}_i$ such that (i) $\tilde{b}_i$ generates an intermediate state distribution between b_i and b_i': $f(\cdot\,|\tilde{b}_i, b_{-\mathrm{i}}) = \lambda f(\cdot\,|b_i, b_{-\mathrm{i}}) + (1-\lambda) f(\cdot\,|b_i', b_{-\mathrm{i}})$, and (ii) the cost of taking $\tilde{b}_i$ is also intermediate between b_i and b_i': $c_i(\tilde{b}_i) = \lambda c_i(b_i) + (1-\lambda) c_i(b_i')$. The hypothesis of Theorem 2 also requires that such intermediate actions are on some deviation paths. Under a such condition, agent i has no incentive to deviate to an action b_i', which is possibly distant from the targeted action b_i, if and only if he has no incentive to deviate to any intermediate actions between b_i and b_i'. Accordingly, under such an assumption, the local optimality condition (with respect to each deviation path) is sufficient for global optimality of the action choice.

Theorems 1 and 2 exhibit the one-to-one tradeoff between the dimensionality of the action space and that of payment rules that induce a targeted action profile. The variety of desirable payment rules sharply decreases as the action space becomes larger in the following sense. Let a mechanism successfully induces a targeted action profile under the current action space. Now, assume that there arises a new class of actions (deviation path). With the new (larger) action space, the mechanism, which performed well under the old (smaller) action space, *generically* fails to induce the targeted action profile because of the reduction of the dimensionality. In this sense, even when the largeness of the action space is intermediate (i.e., there are only a few deviation paths), the inducibility condition imposes a severe restriction on payment rules.

As the number of deviation paths increases, $X_i(b,g)$ shrinks. From now, we turn our eyes into an extreme case. Let

$$T(\Omega) \equiv \left\{ t \in R^{|\Omega|} : \sum_{\omega \in \Omega} t(\omega) = 0 \right\}.$$

Clearly, $\dim T(\Omega) = |\Omega| - 1$. For every $\alpha \in [-1,1] \setminus \{0\}$ and deviation path β_i,

$$\frac{f(\cdot\,|\beta_i(\alpha), b_{-i}) - f(\cdot\,|b)}{\alpha} \in T(\Omega).$$

Since $T(\Omega)$ is a closed set, its limit at $\alpha \to 0$ also belongs to $T(\Omega)$ (i.e., $t(\beta_i) \in T(\Omega)$). Hence, there are at most $|\Omega| - 1$ deviation paths whose tangents are linearly independent. We say that agent i's action space B_i is rich at action profile b if there indeed exist $|\Omega| - 1$ deviation paths whose tangents are linearly independent.

Definition 3: Agent i's action space B_i is *rich at* action profile b if there exist $|\Omega| - 1$ deviation paths of (i,b) whose tangents are linearly independent. An action space B is *rich at* b if B_i is rich at b for all $i \in N$.

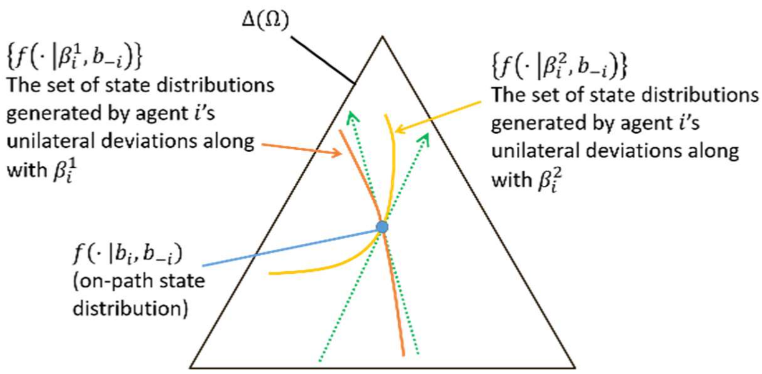

Figure 3: Rich Action Space. The case of $|\Omega| = 3$. B_i is rich at b if there are two $(= |\Omega| - 1)$ different deviation paths of (i,b).

When B_i is rich at b, Theorem 1 implies that $\dim X_i\,(b,g) = 1$. Accordingly, whenever x_i and $\tilde{x}_i$ induce b for agent i, then x_i and $\tilde{x}_i$ are the same up to constants. Hence, we have proved the following equivalence theorem:

Theorem 3: *Suppose that* B_i *is rich at* b *and* (g, x_i) *induces* b *for agent* i. *Then, the following are equivalent:*

(i) $x_i \in X_i(b,g)$ *and* $\tilde{x}_i \in X_i(b,g)$.

(ii) $\tilde{x}_i = x_i + \bar{z}_i$ *for some* $\bar{z}_i \in R$.

Theorem 3 implies the following equivalence properties. Consider an arbitrary combination of an action profile and an allocation rule (b, g). Consider two arbitrary payment rules x and $\tilde{x}$ such that both (g, x) and $(g, \tilde{x})$ induce b. Let $U_i \in R$ and $\widetilde{U}_i \in R$ denote the respective ex-ante expected payoff for each agent $i \in N$:

$$U_i \equiv E[v_i(g(\omega), \omega) - x_i(\omega)|b] - c_i(b_i),$$
$$\widetilde{U}_i \equiv E[v_i(g(\omega), \omega) - \tilde{x}_i(\omega)|b] - c_i(b_i).$$

By Theorem 3, we have the following three equivalence results.

(i) the ex-post payment for each agent i is unique up to constants:

$$\tilde{x}_i(\omega) = x_i(\omega) - U_i + \widetilde{U}_i \text{ for all } \omega \in \Omega.$$

(ii) the ex-post revenue for the central planner is unique up to constants:

$$\sum_{i \in N} \tilde{x}_i(\omega) = \sum_{i \in N} x_i(\omega) - \sum_{i \in N} \left(U_i - \widetilde{U}_i\right) \text{ for all } \omega \in \Omega.$$

(iii) the ex-post payoff for each agent i is unique up to constants: for all $\omega \in \Omega$,

$$v_i(g(\omega), \omega) - \tilde{x}_i(\omega) - c_i(b_i)$$
$$= \{v_i(g(\omega), \omega) - x_i(\omega) - c_i(b_i)\} + U_i - \widetilde{U}_i.$$

5. Hidden Information

This section studies the relationship between the incentives for taking targeted actions and incentives for truthfully reporting private types. First, we define ex-post incentive compatibility in a standard manner.

Definition 4 (Ex-Post Incentive Compatibility): A mechanism (g, x) is *ex-post incentive compatible* (EPIC) if truth-telling is an ex-post equilibrium; for every $i \in N$, $\omega \in \Omega$, and $\widetilde{\omega}_i \in \Omega_i$, we have

$$v_i(g(\omega), \omega) - x_i(\omega) \geq v_i(g(\widetilde{\omega}_i, \omega_{-i}), \omega) - x_i(\widetilde{\omega}_i, \omega_{-i}).$$

By Theorem 3, when the action space is rich, mechanisms that induce the same action profile is identical up to constant. Accordingly, such mechanisms are also equivalent in terms of incentives for reporting. In other words, once we design a payment rule to induce a targeted action profile, we have no degree of freedom for making a payment rule to incentivize truthful reporting.

Corollary 1: *Consider an arbitrary combination of an allocation rule and an action profile* (b, g)*. Suppose that* B *is rich at* b*, and a mechanism* (g, x) *induces* b*. If* (g, x) *is not EPIC, then there is no payment rule* $\tilde{x}$ *such that* $(g, \tilde{x})$ *induces* b *is EPIC.*

Corollary 1 indicates that if there is a mechanism that induces an action profile b but fails to be incentive compatible, then there is no mechanism that satisfies both inducibility and incentive compatibility.

6. Efficiency

This section studies efficient mechanism design (i.e., maximization of the social welfare). We denote the valuation function of the central planner by $v_0 : A \times \Omega \to R$. An allocation rule g is said to be *allocatively efficient* if

$$\sum_{i\in N\cup\{0\}} v_i\,(g(\omega),\omega) \geq \sum_{i\in N\cup\{0\}} v_i\,(a,\omega)$$

for all $a \in A$ and $\omega \in \Omega$. A combination of an action profile and an allocation rule (b,g) is said to be *fully efficient* if g is allocatively efficient and the selection of b maximizes the total welfare in expectation—i.e., it satisfies

$$E\left[\sum_{i\in N\cup\{0\}} v_i(g(\omega),\omega)\middle|\, b\right] - \sum_{i\in N\cup\{0\}} c_i(b_i) \geq E\left[\sum_{i\in N\cup\{0\}} v_i(g(\omega),\omega)\middle|\, \tilde{b}\right] - \sum_{i\in N\cup\{0\}} c_i(\tilde{b}_i)$$

for all $\tilde{b} \in B$. A mechanism (g,x) is *fully efficient* if there exists an action profile $b \in B$ such that (g,x) induces b and (b,g) is fully efficient.[10]

A payment rule x is said to be *VCG* if there exists $y_i : \Omega_{-i} \to R$ for each $i \in N$ such that

$$x_i(\omega) = -\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\,(g(\omega),\omega) + y_i(\omega_{-i})$$

for all $i \in N$ and $\omega \in \Omega$. A VCG payment rule gives each agent $i \in N$ the monetary amount equivalent to the other agents' welfare plus the central planner's welfare (i.e., $\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\,(g(\omega),\omega)$) and retrieves the monetary payment $y_i(\omega_{-i})$ that is independent of ω_i.

We now introduce a subclass of VCG mechanisms, the pure-VCG mechanism. A payment rule x is said to be *pure-VCG* if it is VCG and $y_i(\cdot)$ is constant for each $i \in N$; there exists a vector $\bar{y} = (\bar{y}_i)_{i\in N} \in R^n$ such that

$$x_i(\omega) = -\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\,(g(\omega),\omega) + \bar{y}_i$$

for all $i \in N$ and $\omega \in \Omega$. Pure-VCG payment rules are special cases of VCG payment rules, where the central planner imposes on each agent i a fixed amount $\bar{y}_i$ as a non-incentive term. We term a combination of efficient allocation rule and a VCG payment rule (pure-VCG payment rule) a *VCG mechanism* (*pure-VCG mechanism*).

[10] Our definition of full efficiency of mechanisms is different from the efficiency defined in Hatfield, Kojima, and Kominers (2015) in that we fix agents' action spaces and cost functions.

The pure-VCG mechanism can be interpreted as a generalization of the selling-out contract of the principal–agent problems. The selling-out contract perfectly aligns the output and the agent's payoff (by selling out the firm from the principal to the agent). Under the selling-out contract, the agent's problem becomes isomorphic to the maximization of the expected profit; thus, the first-best effort level is induced (Harris and Raviv, 1979).

In a general efficient mechanism design problem, the "output" is the social welfare, and the pure-VCG mechanism perfectly aligns each agent's payoff with the social welfare. Accordingly, pure-VCG induces efficient actions. We further show that, when the action space is rich, only pure-VCG mechanism can be fully efficient: all the other mechanisms are taken advantage of by some actions.

Theorem 4:

(i) *Pure-VCG mechanisms are fully efficient.*

(ii) *Suppose that, for every fully efficient* (b, g), B *is rich at* b. *Then, a mechanism is fully efficient if and only if it is pure-VCG.*

Proof: If (b, g) is fully efficient and x is pure-VCG, then, for every $i \in N$ and $b_i' \in B_i$,

$$\begin{aligned}
&E[v_i(g(\omega),\omega) - x_i(\omega)|b] - c_i(b_i) - E[v_i(g(\omega),\omega) - x_i(\omega)|b_i', b_{-i}] + c_i(b_i') \\
&= E\left[\sum_{j \in N \cup \{0\}} v_j(g(\omega),\omega)\middle| b\right] - c_i(b_i) - \left[\sum_{j \in N \cup \{0\}} v_j(g(\omega),\omega)\middle| b_i', b_{-i}\right] + c_i(b_i') \\
&\geq 0,
\end{aligned}$$

which implies that (g, x) induces b. Furthermore, if B is rich at b, it is clear from Theorem 3 and the definition of the pure-VCG payment rule that only pure-VCG mechanisms will induce b.

Q.E.D.

Note that any pure-VCG mechanism can be constructed without utilizing any detailed information about (f, B, c). In this sense, pure-VCG follows the Wilson doctrine (Wilson, 1987): it is free from the probability assessment. When the action space is rich, even if such information is available, the pure-VCG mechanism is the only choice for maximizing the social welfare.

We say that we have *private values* if, for every $i \in N$, (i) the valuation $v_i(a,\omega)$ is independent of the other agents' type profile $\omega_{-0-i} \equiv (\omega_j)_{j\in N\setminus\{i\}} \in \Omega_{-0-i} \equiv \times_{j\in N\setminus\{i\}} \Omega_j$, and (ii) $v_0(a,\omega)$ is independent of ω_{-0}. We write $v_i(a,\omega_0,\omega_i)$ instead of $v_i(a,\omega)$ for each $i \in N$, and $v_0(a,\omega_0)$ instead of $v_0(a,\omega)$.[11] It is well-known that, when we have private values, VCG mechanism satisfies EPIC.[12] Accordingly, the pure-VCG mechanism also solves the hidden information problem.

Theorem 4 also implies that achieving full efficiency is generally impossible, when we have interdependent values (i.e., the private-value assumption is violated). Even with interdependent values, it is possible, under certain assumptions, to achieve the allocative efficiency (using some mechanisms other than VCG).[13] It is also well-known that VCG mechanisms rarely satisfy EPIC (or a weaker incentive compatibility condition) with interdependent values. On the other hand, Theorem 4 indicates that the pure-VCG mechanism is the only choice for inducing an efficient action profile. Hence, with interdependent values, (i) an efficient action profile, (ii) an efficient allocation rule, and (iii) incentive compatibility for reporting cannot be achieved at the same time, typically.[14] This generalizes the observation for the mechanism design problem with information acquisition, studied by Bergemann and Välimäki (2002).

Although any VCG mechanism is incentive compatible and allocatively efficient in the private-value case, they do not always induce an efficient action profile. For example, consider the *pivot mechanism* (g,x), which is a VCG mechanism specified by

$$x_i(\omega) = -\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\big(g(\omega),\omega_0,\omega_j\big) + \min_{\omega_i'\in\Omega_i} \sum_{j\in N\cup\{0\}} v_j\big(g(\omega_i',\omega_{-i}),\omega_0,\omega_j\big).$$

[11] We permit the valuations to depend on the public signal ω_0. We write $v_0(a,\omega_0,\omega_0)$ instead of $v_0(a,\omega_0)$.

[12] It is well-known that, with the assumption of private values, EPIC is equivalent to *incentive compatibility in dominant strategies* (DIC); for every $i \in N$, $\omega \in \Omega$, and $\tilde{\omega}_i \in \Omega_i$, we have

$$v_i(g(\omega),\omega_0,\omega_i) - x_i(\omega) \geq v_i(g(\tilde{\omega}_i,\omega_{-i}),\omega_0,\omega_i) - x_i(\tilde{\omega}_i,\omega_{-i}).$$

[13] See, for example, Crémer and McLean (1985, 1988), Maskin (1992), Dasgupta and Maskin (2000), Jehiel and Moldovanu (2001), and Bergemann and Välimäki (2002).

[14] Mezzetti (2004) shows that, if the realized valuation $v_i(g(\omega),\omega)$ is observable as an ex-post public signal and is contractible, even with interdependent values, (allocatively) efficient mechanisms can be implemented providing any VCG expected payoffs to each agent (see Noda, 2019 for more general ex-post signals). If these schemes are available, full efficiency is achievable even with interdependent values.

Clearly, the pivot mechanism is not pure-VCG because the "non-incentive term" is not constant. The following proposition demonstrates a sufficient condition under which the pivot mechanism fails to achieve full efficiency:

Proposition 2: *Suppose that allocation rule* g *is allocatively efficient. Define* $z_i : \Omega \to R$ *by*

$$z_i(\omega) = \min_{\omega_i' \in \Omega_i} \sum_{j \in N \cup \{0\}} v_i\big(g(\omega_i', \omega_{-i}), \omega_0, \omega_j\big). \tag{8}$$

Then, a pivot mechanism fails to achieve full efficiency if, for every $b \in B$ *such that* (b, g) *is fully efficient, there exists an agent* $i \in N$ *and a deviation path* β_i *of* (i, b) *such that*

$$z_i \cdot t(\beta_i) \neq 0.$$

Proof: z_i, defined by (8), is the difference between a pure-VCG payment rule and the pivot payment rule. Since pure-VCG mechanisms induce b, it follows from Proposition 1 that the pivot mechanism induces b for i only if $z_i \cdot t(\beta_i) = 0$.

Q.E.D.

Proposition 2 indicates that the pivot mechanism induces an efficient action profile only when a special congruence condition is satisfied. We have no reasons to expect that every tangent of deviation path will be perpendicular to z_i, which is defined by (8). Therefore, even if there exists at least *one* deviation path around the efficient action profile, the pivot mechanism *generically fails* to achieve full efficiency.[15]

Remark 2: In Appendix B, we propose an open-bidding procedure that implements pure-VCG mechanisms. In our procedure, the mechanism determines the compensation for losers, rather than the payment to the winner (as in standard auctions). The mechanism starts with a sufficiently high price and gradually descends the price until one bidder declares to win the object. The winner

[15] Prior to our work, Krähmer and Strausz (2007) showed that the pivot mechanism generically fails to induce an efficient level of investment when each agent takes one-dimensional ex-ante investments. Proposition 2 generalizes this observation.

obtains nothing, and the losers are compensated with the price determined by the mechanism. This procedure is in contrast to the ascending (English) auction implementing the pivot mechanism. For further details, the reader is directed to Appendix B.

Remark 3: Theorem 4 and Proposition 2 indicate that when agents can take hidden actions, it is difficult to earn non-negative revenue while achieving full efficiency. Although the pure-VCG mechanism achieves full efficiency, it cannot earn non-negative revenue in a very general class of environment. For further details, the reader is directed to Appendix C.

7. Independent Types and Interim Equivalence

Thus far, we have assumed that each agent's action choice (possibly) have externality effects (i.e., may influence the distribution of the entire state ω, rather than his private type ω_i). We have seen that, even if the dimensionality of unilateral deviations is low, the pivot mechanism generally fails to achieve full efficiency (Proposition 2). If agents can take various unilateral deviations (i.e., if the action space is rich), only pure-VCG mechanism can be full efficient (Theorem 4).

This section excludes externality effects by assuming that one agent's action only influences his own type distribution. Even with such an assumption, Theorem 1 holds without any modifications. Parallel to Theorem 4, we can also derive the equivalence theorem from Theorem 1. However, the equivalence result is interim (rather than ex-post), and therefore, the set of mechanisms inducing a targeted action profile becomes significantly larger.

As in Bergemann and Välimäki (2002) and Hatfield, Kojima, and Kominers (2015), this section assumes independent types of information structure. We have *independent types* if each agent i's action choice $b_i \in B_i$ influences only the marginal distribution of the agent's own type ω_i; for every $\omega \in \Omega$ and $b \in B$, we have

$$f(\omega|b) = f_0(\omega_0)\prod_{i\in N} f_i(\omega_i|b_i).$$

Here, $f_i(\cdot\,|b_i)$ denotes the marginal distribution of each agent i's type ω_i, which is assumed to depend only on b_i, and $f_0(\cdot)$ denotes the distribution of the public signal ω_0, which is assumed to be independent of the action profile, $b \in B$. When we have independent types, agent i's action space is equivalent to the set of available marginal distributions on agent i's type. With independent types, for every $\xi : \Omega \rightarrow R$ and $\xi_i : \Omega_i \rightarrow R$, we can simply write $E_{\omega_{-i}}[\xi(\tilde{\omega}_i, \omega_{-i})|b_{-i}]$ and $E[\xi_i(\omega_i)|b_i]$ instead of $E_{\omega_{-i}}[\xi(\tilde{\omega}_i, \omega_{-i})|b, \omega_i]$ and $E[\xi_i(\omega_i)|b]$, respectively.

Even with independent types, Theorem 1 remains true without any modification; thus, there is a one-to-one tradeoff between the number of deviation paths and the dimensionality of payment rules with which a targeted action profile is induced. However, with independent types, the action set cannot be rich (i.e., cannot have $|\Omega| - 1$ distinct deviation paths). With independent types, the choice of b_i is isomorphic to the choice of $f_i \in \Delta(\Omega_i)$, where $\dim \Delta(\Omega_i) = |\Omega_i|$. Accordingly, there are at most $|\Omega_i| - 1$ deviation paths whose tangents are linearly independent. We say that the action space B_i is individually rich at b_i if such deviation paths actually exist.

Definition 5 (Individual Richness): Agent i's action space B_i is *individually rich at* b_i if there exist $|\Omega_i| - 1$ deviation paths of (i, b) whose tangents are linearly independent.

When agent i's action space is individually rich, agent i can manipulate his own marginal type distribution to $|\Omega_i| - 1$ different directions. We cannot find a larger number of deviation paths whose tangents are linearly independent, as long as we maintain the independent-type assumption.

Parallel to Theorem 3, when the action space is individually rich, the payment rule inducing a targeted action profile become unique up to "constants" in the interim sense.

Theorem 5: *Suppose that* B_i *is individually rich at* b *and* (g, x_i) *induces* b *for agent* i*. Then, the following are equivalent:*

(i) $x_i \in X_i(b, g)$ *and* $\tilde{x}_i \in X_i(b, g)$.

(ii) $\tilde{x}_i = x_i + r_i$ *for some* $r_i : \Omega \rightarrow R$ such that $E_{\omega_{-i}}[r_i(\omega_i, \omega_{-i})|b_{-i}]$ is independent of $\omega_i \in \Omega_i$.

Proof: See Appendix A.

The distribution of ω_{-i} is determined solely by the choice of b_{-i}, which agent i cannot manipulate. Accordingly, agent i's unilateral deviation does not change the value of $E_{\omega_{-i}}[r_i(\omega_i, \omega_{-i})|b_{-i}]$ whenever it is independent of ω_i. Hence, the interim equivalence is sufficient. When the action space is individually rich, agent i can manipulate the distribution of ω_i in full-dimensional directions. Hence, if $E_{\omega_{-i}}[r_i(\omega_i, \omega_{-i})|b_{-i}]$ depends on ω_i, agent i has a way to take advantage of it. Accordingly, the interim equivalence is necessary.

Now, we consider fully efficient mechanisms. A payment rule x is said to be *expectation-VCG* if, for each $i \in N$, there exist $r_i : \Omega \to R$ such that for every $i \in N$ and $\omega \in \Omega$, we have

$$x_i(\omega) = -\sum_{j \in N \cup \{0\} \backslash \{i\}} v_j(g(\omega), \omega) + r_i(\omega),$$

and $E_{\omega_{-i}}[r_i(\omega_i, \omega_{-i})|b_{-i}]$ is independent of $\omega_i \in \Omega_i$. We term the combination of an efficient allocation rule and expectation-VCG payment rule an *expectation-VCG mechanism*. Clearly, any VCG mechanism (including pure-VCG) belongs to expectation-VCG.

It is straightforward to show that any expectation-VCG mechanism can induce an efficient action profile (Rogerson, 1992). Furthermore, by Theorem 5, any mechanism inducing the efficient action profile is interim equivalent if the action space is individually rich. Accordingly, only expectation-VCG can be fully efficient.

Theorem 6: *Under the assumption of independent types:*

(i) *Any expectation-VCG mechanism is fully efficient.*

(ii) *Suppose that, for every fully efficient* (b, g), b *is individually rich. Then, a mechanism is fully efficient if and only if it is expectation-VCG.*

Proof: See Appendix A.

Trivially, expectation-VCG mechanism does not satisfy EPIC in general (even with private values). On the other hand, it is straightforward to see that when the type distribution is fixed and we have private values, expectation-VCG mechanism is Bayesian incentive compatible (BIC): if a profile of truthful reporting consists a Bayesian Nash equilibrium.

In our model, standard BIC is not sufficient for preventing deviations. The traditional definition of BIC only considers deviations in reporting (assuming a type distribution to be fixed). However, in our model, agents can make a joint deviation in actions and reporting. To consider an incentive for joint deviations, we define the following condition:[16]

Definition 5 (Bayesian Inducibility and Incentive Compatibility): A combination of an action profile and a mechanism $(b,(g,x))$ is *Bayesian inducible and incentive compatible* (BIIC) if the selection of action profile b at stage 2 and the truthful revelation at stage 4 result in a perfect Bayesian equilibrium; for every agent $i \in N$, action $b_i' \in B_i$, and reporting strategy $\sigma_i : \Omega_i \to \Omega_i$,

$$\begin{aligned}&E[v_i(g(\omega),\omega) - x_i(\omega)|b] - c_i(b_i)\\ &\geq E[v_i(g(\sigma_i(\omega_i),\omega_{-i}),\omega_{-i}) - x_i(\sigma_i(\omega_i),\omega_{-i})|b_i',b_{-i}] - c_i(b_i').\end{aligned}$$

The expectation-VCG mechanism satisfies not only a traditional BIC but also BIIC.

Proposition 3: *Suppose that we have private values and independent types. Then, expectation-VCG mechanism satisfies BIIC.*

Proof: See Appendix A.

For example, AGV mechanism, established by Arrow (1979) and D'Aspremont and Gérard-Varet (1979), is an expectation-VCG mechanism specified by

$$\begin{aligned}r_i(\omega) = &\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\big(g(\omega),\omega_j\big) - E_{\widetilde{\omega}_{-i}}\left[\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\big(g(\omega_i,\widetilde{\omega}_{-i}),\widetilde{\omega}_j\big)\middle| b_{-i}\right]\\ &+\frac{1}{n-1}\sum_{j\in N\setminus\{i\}} E_{\widetilde{\omega}_{-i}}\left[\sum_{h\in N\cup\{0\}\setminus\{j\}} v_h\big(g\big(\omega_j,\widetilde{\omega}_{-j}\big),\widetilde{\omega}_h\big)\middle| b_{-i}\right].\end{aligned}$$

[16] When we consider EPIC, the incentive condition for actions (inducibility) and reporting (EPIC) can be separated. This is because the definition of EPIC is independent of the underlying type distribution. However, the definition of the Bayesian incentive compatibility is dependent on the type distribution, and therefore, it cannot be defined separately from the inducibility condition.

By definition, the AGV mechanism satisfies *budget balance*: for all $\omega \in \Omega$, $\sum_{i \in N} x_i(\omega) = 0$. Accordingly, with independent types the central planner can achieve full efficiency along with BIIC and budget balance.

The observation in this section articulates our contribution relative to previous works. Rogerson (1992) proved that, with independent and private values, expectation-VCG mechanisms induce efficient action profiles. Bergemann and Välimäki (2002) studied information acquisition for each agent's own hidden state with independent types and private values and demonstrated that VCG mechanisms achieve full efficiency. We derive the full characterization while allowing general action spaces. Hatfield, Kojima, and Kominers (2015) showed that when detailed information about the environment (action space, cost function, and state distribution associated with each action profile) is not available, the VCG mechanism is the unique fully efficient mechanism.[17] This result is a special case of our Theorem 6 in the sense that only VCG is "detail-free" (i.e., it can be constructed without using information about (f, B, c)) among all expectation-VCG mechanisms.

8. Conclusion

We have studied a general mechanism design problem in which agents take hidden actions that determine the state distribution. The variety of action choices substantially restricts the class of mechanisms that successfully induce a desired action profile. Indeed, the popular pivot mechanism generically fails to induce an efficient action profile, if there exists at least one class of actions (deviation path) an agent may take. If agents can take various unilateral deviations (i.e., if the action space is rich at the targeted action profile), ex-post payoffs, ex-post payments, and ex-post revenues are all unique up to constants. In particular, only pure-VCG mechanisms are able to achieve full efficiency. The pure-VCG mechanism has a very restrictive form that is not compatible with many demanded properties (e.g., ex-post individual rationality, non-negative revenue, budget balance). Hence, this result could be interpreted as a negative result.

In order to break down this impossibility result, the central planner must collect detailed information about the set of possible action choices. If the central planner is not fully aware of the

[17] On the other hand, Tomoeda (2017) pointed out that the VCG mechanism fails to induce an efficient action profile as a *unique* equilibrium even under the absence of externality effects.

set of all possible actions, then she must use a robust and detail-free mechanism, such as pure-VCG, that achieves full efficiency regardless of the detailed structure of actions and state distributions. However, if the central planner knows the action set, she can design a detail-dependent mechanism that performs well for a particular problem she faces, unless the action space is actually rich. Even if the action space is rich, agents may pay limited attention to action choices as the wide availability of actions make it difficult to find exactly optimal actions. If the planner finds that agents dismiss some dimensions of actions, she could also exclude such actions from the design problem. The detailed knowledge about the set of actions, state distributions associated with each action profile, and agents' perception is necessary for these two countermeasures.

Alternatively, the central planner can also decide to compromise full efficiency to enhance the degree of freedom of the mechanism design. On the one hand, the planner can use the pivot mechanism to achieve allocation efficiency and non-negative revenue, while compromising on action efficiency. On the other hand, the planner can also offer reward or punishment to prevent action deviations while compromising on allocation efficiency. It will be interesting to consider the optimal way to compromise the action and allocation efficiencies; however, since it is beyond the scope of this paper, this question should be answered in future research.

Appendix A: Omitted Proofs

Proof of Theorem 5: (i) $\Rightarrow$ (ii) is straightforward from the fact that $E_{\omega_{-i}}[r_i(\omega_i, \omega_{-i})|b_{-i}]$ is independent of ω_i and b_i.

To show the other direction, we first prove the following lemma.

Lemma 1: With the assumption of independent types, if β_i is a deviation path of b and $t(\beta_i) \in R^{|\Omega|}$ is its tangent, then there exists $t_i(\beta_i) \in R^{|\Omega_i|}$ such that

$$\lim_{\alpha \to 0} \frac{f_i(\cdot\,|\beta_i(\alpha)) - f_i(\cdot\,|b_i)}{\alpha} \equiv t_i(\beta_i).$$

Furthermore, for all deviation paths $\beta_i^1, \cdots, \beta_i^K$, their respective tangents $t(\beta_i^1)$, …, $t(\beta_i^K)$ are linearly independent if and only if $t_i(\beta_i^1)$, …, $t_i(\beta_i^K)$ are linearly independent.

Proof: By independent types,

$$f(\omega|\beta_i(\alpha), b_{-i}) - f(\omega|b_i, b_{-i}) = \prod_{j\neq i} f_j(\omega_j|b_j) \cdot \big(f_i(\omega_i|\beta_i(\alpha)) - f_i(\omega_i|b_i)\big).$$

Accordingly, we have

$$\text{(A1)} \qquad \begin{aligned} t(\omega, \beta_i) &= \prod_{j\neq i} f_j(\omega_j|b_j) \cdot \lim_{\alpha\to 0} \frac{f_i(\omega_i|\beta_i(\alpha)) - f_i(\omega_i|b_i)}{\alpha} \\ &= \prod_{j\neq i} f_j(\omega_j|b_j) \cdot t_i(\omega_i, \beta_i). \end{aligned}$$

This indicates that whenever $t(\beta_i)$ exists, $t_i(\beta_i)$ also exists; thus, $t_i(\beta_i)$ is well-defined.

Suppose that $t(\beta_i^1)$, …, $t(\beta_i^K)$ are not linearly independent; then, there exists $(\lambda_k)_{k=0}^K$ such that $\lambda_k \neq 0$ for some $k \in \{1, \cdots, K\}$, and

$$\sum_{k=1}^{K} \lambda_k \cdot t(\omega, \beta_i^k) = 0$$

for all $\omega \in \Omega$. Fix some $\omega_{-i} \in \Omega_{-i}$ arbitrarily. It follows from (A1) that

$$\sum_{k=1}^{K} \lambda_k \cdot t_i(\omega_i, \beta_i^k) = \frac{1}{\prod_{j\neq i} f_j(\omega_j|b_j)} \sum_{k=1}^{K} \lambda_k \cdot t(\omega, \beta_i^k) = 0$$

for all $\omega_i \in \Omega_i$. Hence, $t_i(\beta_i^1)$, …, $t_i(\beta_i^K)$ are not linearly independent either. If $t_i(\beta_i^1)$, …, $t_i(\beta_i^K)$ are not linearly independent, we can show that $t(\beta_i^1)$, …, $t(\beta_i^K)$ are not independent in a similar manner.

Q.E.D.

Proof of Theorem 5 (continued): Suppose that both (g, x) and $(g, \tilde{x})$ induces b. Then, for all $i \in N$, the following first-order condition of necessary condition must be satisfied: for $k = 1, \cdots, |\Omega_i| - 1$,

$$\frac{\partial}{\partial \alpha}\left\{E\left[v_i(g(\omega),\omega) - x_i(\omega)|\beta_i^k(\alpha), b_{-i}\right] - c_i\left(\beta_i^k(\alpha)\right)\right\}\Big|_{\alpha=0} = 0, \text{and}$$

$$\frac{\partial}{\partial \alpha}\left\{E\left[v_i(g(\omega),\omega) - \tilde{x}_i(\omega)|\beta_i^k(\alpha), b_{-i}\right] - c_i\left(\beta_i^k(\alpha)\right)\right\}\Big|_{\alpha=0} = 0.$$

Comparing these equations, we obtain

(A2) $$\frac{\partial}{\partial \alpha}E\left[x_i(\omega) - \tilde{x}_i(\omega)|\beta_i^k(\alpha), b_{-i}\right]\Big|_{\alpha=0} = 0$$

for $k = 1,\cdots,|\Omega_i| - 1$. With independent types, (A2) can be rewritten as

(A3) $$\frac{\partial}{\partial \alpha}E_{\omega_i}\left[E_{\omega_{-i}}[x_i(\omega_i,\omega_{-i}) - \tilde{x}_i(\omega_i,\omega_{-i})|b_{-i}]|\beta_i^k(\alpha)\right]\Big|_{\alpha=0} = 0$$

for $k = 1,\cdots,|\Omega_i| - 1$. Define

$$w_i(\omega_i) \equiv E_{\omega_{-i}}[x_i(\omega_i,\omega_{-i}) - \tilde{x}_i(\omega_i,\omega_{-i})|b_{-i}].$$

Then, (A3) is equivalent to

(A4) $$w_i \cdot t_i(\beta_i^k) = 0$$

for $k = 1,\cdots,|\Omega_i| - 1$, where $t_i(\beta_i^1)$, …, $t_i(\beta_i^{|\Omega_i|-1})$ are obtained by Lemma 1. Since $\{t_i(\beta_i^k)\}_{k=1}^{|\Omega_i|-1}$ are linearly independent,

$$\dim\left\{w_i \in R^{|\Omega_i|} : w_i \cdot t_i(\beta_i^k) = 0 \text{ for } k = 1,\cdots,|\Omega_i| - 1\right\} = 1.$$

This and (A4) imply that there exists $\bar{z}_i \in R$ such that

$$w_i(\omega_i) = E_{\omega_{-i}}[x_i(\omega_i,\omega_{-i}) - \tilde{x}_i(\omega_i,\omega_{-i})|b_{-i}] = \bar{z}_i$$

for all $\omega_i \in \Omega_i$. Since $r_i(\omega) = x_i(\omega) - \tilde{x}_i(\omega)$, Theorem 5 is proved.

Q.E.D.

Proof of Theorem 6: Whenever (b,g) is fully efficient and x is pure-VCG, then (g,x) induces b. Furthermore, if x is a pure-VCG payment rule, $\tilde{x}$ satisfies $\tilde{x}_i(\omega) = x_i(\omega) + r_i(\omega)$

and $E_{\omega_{-i}}[r_i(\omega_i, \omega_{-i})|b_{-i}]$ is independent of $\omega_i \in \Omega_i$, then $\tilde{x}$ is an expectation-VCG payment rule. Accordingly, by Theorem 5, $(g, \tilde{x})$ induces b if and only if $\tilde{x}$ is an expectation-VCG payment rule.

Q.E.D.

Appendix B: Open-Bid Implementation of Pure-VCG

In a single-unit auction with private values, we can implement a pure-VCG mechanism with an open-bid procedure, just as we can implement a pivot mechanism with an ascending auction. The allocation $a \in A \equiv \{1,2,\cdots,n\}$ specifies the agent who is the winner (who obtains the object). Each agent's type $\omega_i \in \Omega_i \subset R_+$ specifies the payoff of that agent when she obtains the object, i.e.,

$$v_i(a, \omega_i) = \begin{cases} \omega_i & \text{if } a = i, \\ 0 & \text{otherwise.} \end{cases}$$

An efficient allocation rule g is given by

$$g(\omega) \in \underset{i \in N}{\operatorname{argmax}}\, \omega_i \text{ for all } \omega \in \Omega.$$

The pure-VCG payment rule is given by

$$x_i(\omega) = \begin{cases} \bar{y}_i & \text{if } g(\omega) = i, \\ -\omega_{g(\omega)} + \bar{y}_i & \text{otherwise,} \end{cases}$$

where $\bar{y}_i \in R$ is an arbitrary real number for each $i \in N$.

We consider the following open-bid descending procedure:

1. Each agent pays a fixed participation fee $\bar{y}_i$ to the central planner at the beginning.
2. The mechanism initially sets the price sufficiently high, and then gradually descends the price.
3. When an agent declares taking the object, this descending procedure is immediately terminated. The agent who declares taking the object becomes the winner and obtains this object.

4. The central planner does *not* require the winner to pay any additional fee. Instead, the central planner gives any other agent (i.e., any loser) the price at the ending time as a compensation.

It is a weakly dominant strategy[18] for each agent to "declare taking the object" exactly when the current price is equal to the agents' value. If all the agents adopt this dominant strategy, then (i) the agent with the highest value wins, (ii) the winner's payment is $\bar{y}_i$, and (iii) the loser's payment is $-\omega_{g(\omega)} + \bar{y}_i$. Accordingly, this open-bid descending procedure generates the same outcome as a pure-VCG mechanism.

The pivot mechanism (the second-price auction) is implemented by the popular ascending auction that determines the winner's payment. By contrast, our procedure descends the price and determines the compensation for losers, implementing a pure-VCG mechanism.

Appendix C: Individual Rationality and Revenues

Assuming private values, this section studies whether the central planner can achieve efficiency *without deficits*. We define the central planner's ex-post revenue by

(C1) $$v_0(g(\omega), \omega_0) + \sum_{i \in N} x_i(\omega),$$ [19]

and the expected revenue in the ex-ante term by

$$E\left[v_0(g(\omega), \omega_0) + \sum_{i \in N} x_i(\omega) \middle| b \right].$$

We introduce three notions of individual rationality below. The timings of the exit opportunities are depicted in Figure 1.

---

[18] This open-bid descending procedure satisfies *obvious strategy-proofness* as defined in Li (2017), same as the ascending auctions.

[19] Note that the central planner's revenue includes not only payments from the agents but also the valuation of the central planner. In the single agent case, where the allocation space is degenerate, the value of (C1) corresponds to the principal's payoff in a standard principal–agent model with a risk-neutral principal and agent.

Definition 6 (Ex-ante Individual Rationality): A combination of an action profile and a mechanism $(b,(g,x))$ satisfies *ex-ante individual rationality (hereinafter EAIR)* if for all $i \in N$, we have

$$E[v_i(g(\omega),\omega_0,\omega_i) - x_i(\omega)|b] - c_i(b_i) \geq 0$$

Definition 7 (Interim Individual Rationality): A combination of an action profile and a mechanism $(b,(g,x))$ satisfies *interim individual rationality* (IIR) if

$$E_{\omega_{-i}}[v_i(g(\omega_i,\omega_{-i}),\omega_0,\omega_i) - x_i(\omega_i,\omega_{-i})|b,\omega_i] \geq 0$$

for all $i \in N$ and $\omega_i \in \Omega_i$, where

$$E_{\omega_{-i}}[\xi(\widetilde{\omega}_i,\omega_{-i})|b,\omega_i] \equiv \sum_{\omega_{-i}\in\Omega_{-i}} \xi(\widetilde{\omega}_i,\omega_{-i})f_{-i}(\omega_{-i}|b,\omega_i)$$

denotes the expectation of a function $\xi(\widetilde{\omega}_i,\cdot):\Omega_{-i} \to R$ conditional on (b,ω_i).

Definition 8 (Ex-post Individual Rationality): A mechanism (g,x) satisfies the *ex-post individual rationality* (EPIR) if for all $i \in N$ and $\omega \in \Omega$, we have

$$E_{\omega_{-i}}[v_i(g(\omega_i,\omega_{-i}),\omega_0,\omega_i) - x_i(\omega_i,\omega_{-i})|b,\omega_i] \geq 0.$$

EPIR implies IIR; however, IIR does not necessarily imply EAIR because the cost for the action choice at stage 2 is sunk. The following proposition shows that EPIR implies EAIR:

Proposition 3: *Suppose that (g,x) induces b. Whenever (g,x) satisfies EPIR, $(b,(g,x))$ satisfies EAIR.*

Proof: Because $c_i(b_i^0) = 0$, it follows from EPIR and inducibility that

$$E[v_i(g(\omega),\omega_0,\omega_i)-x_i(\omega)|b]-c_i(b_i)$$
$$\geq E[v_i(g(\omega),\omega_0,\omega_i)-x_i(\omega)|b_i^0,b_{-i}]-c_i(b_i^0)\geq 0,$$

which implies EAIR.

Q.E.D.

The following proposition calculates the maximal expected revenues (i.e., the least upper-bounds of the central planner's expected revenues):

Proposition 4: *Suppose that* (b,g) *is fully efficient,* B *is rich at* b, *and we have private values. Then, the maximal expected revenue from* $(b,(g,x))$ *in terms of* $x\in\times_{i\in N}X_i(b,g)$ *that satisfies EPIR, IIR, EAIR, and IIR and EAIR are given by*

(C2) $$R^{EPIR}\equiv n\cdot\min_{\omega\in\Omega}\sum_{j\in N\cup\{0\}}v_j\left(g(\omega),\omega_0,\omega_j\right)-(n-1)E\left[\sum_{j\in N\cup\{0\}}v_j\left(g(\omega),\omega_0,\omega_j\right)\middle|\, b\right],$$

(C3) $$R^{IIR}\equiv\sum_{i\in N}\min_{\omega_i\in\Omega_i}E_{\omega_{-i}}\left[\sum_{j\in N\cup\{0\}}v_j\left(g(\omega_i,\omega_{-i}),\omega_0,\omega_j\right)\middle|\, b,\omega_i\right]$$
$$-(n-1)E\left[\sum_{j\in N\cup\{0\}}\left(g(\omega),\omega_0,\omega_j\right)\middle|\, b\right],$$

(C4) $$R^{EAIR}\equiv E\left[\sum_{j\in N\cup\{0\}}v_j\left(g(\omega),\omega_0,\omega_j\right)\middle|\, b\right]-\sum_{j\in N}c_j(b_j),$$

$$R^{EAIR,IIR} \equiv E[v_0(g(\omega),\omega_0)|b] + \sum_{i\in N} \min\left\{ E[v_i(g(\omega),\omega_0,\omega_i)|b] - c_i(b_i), \min_{\omega_i\in\Omega_i} E_{\omega_{-i}}\left[\sum_{j\in N\cup\{0\}} v_j\left(g(\omega_i,\omega_{-i}),\omega_0,\omega_j\right)\middle| b,\omega_i\right] - E\left[\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\left(g(\omega),\omega_0,\omega_j\right)\middle| b\right]\right\}, \tag{C5}$$

respectively.

Proof: By Theorem 4, we can focus on pure-VCG payment rules, where the constraint for EPIR is equivalent to

$$\min_{\omega\in\Omega} \sum_{j\in N\cup\{0\}} v_j\left(g(\omega),\omega_0,\omega_j\right) \geq z_i \text{ for all } i \in N. \tag{C6}$$

We can maximize the expected revenue by letting z_i satisfy (C6) with equality for each $i \in N$. Accordingly, the central planner can receive from each agent i the expected value given by

$$\min_{\omega\in\Omega} \sum_{j\in N\cup\{0\}} v_j\left(g(\omega),\omega_0,\omega_j\right) - E\left[\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\left(g(\omega),\omega_0,\omega_j\right)\middle| b\right],$$

which implies (C2), where we add $E[v_0(g(\omega),\omega_0)|b]$.

Similarly, the constraint for IIR is equivalent to

$$\min_{\omega_i\in\Omega_i} E_{\omega_{-i}}\left[\sum_{j\in N\cup\{0\}} v_j\left(g(\omega_i,\omega_{-i}),\omega_0,\omega_j\right)\middle| \omega_i, b\right] \geq z_i \tag{C7}$$

for all $i \in N$ and $\omega_i \in \Omega_i$. We can maximize the expected revenue by letting z_i satisfy (C7) with equality for each $i \in N$. Accordingly, the central planner can receive from each agent i the expected value given by

$$\min_{\omega_i \in \Omega_i} E_{\omega_{-i}}\left[\sum_{j \in N \cup \{0\}} v_j\left(g(\omega_i, \omega_{-i}), \omega_0, \omega_j\right) \middle| \omega_i, b\right] - E\left[\sum_{j \in N \cup \{0\} \setminus \{i\}} v_j\left(g(\omega), \omega_0, \omega_j\right) \middle| b\right],$$

which implies (C3), where we add $E[v_0(g(\omega), \omega_0)|b]$.

The constraint for EAIR is equivalent to

$$E\left[\sum_{j \in N \cup \{0\}} v_j\left(g(\omega), \omega_0, \omega_j\right) \middle| b\right] - c_i(b_i) \geq z_i \tag{C8}$$

for all $i \in N$. We can maximize the expected revenue by letting z_i satisfy (C8) with equality for each $i \in N$. Accordingly, the central planner can receive from each agent i the expected value given by

$$\begin{aligned} &E\left[\sum_{j \in N \cup \{0\}} v_j\left(g(\omega), \omega_0, \omega_j\right) \middle| b\right] - c_i(b_i) - E\left[\sum_{j \in N \cup \{0\} \setminus \{i\}} v_j\left(g(\omega), \omega_0, \omega_j\right) \middle| b\right] \\ &= E[v_i(g(\omega), \omega_0, \omega_i)|b] - c_i(b_i), \end{aligned}$$

which implies (C4), where we add $E[v_0(g(\omega), \omega_0)|b]$.

The constraint for IIR and EAIR is equivalent to

$$\begin{aligned} \min\Bigg\{ &E\left[\sum_{j \in N \cup \{0\}} v_j\left(g(\omega), \omega_0, \omega_j\right) \middle| b\right] - c_i(b_i), \\ &\min_{\omega_i \in \Omega_i} E_{\omega_{-i}}\left[\sum_{j \in N \cup \{0\}} v_j\left(g(\omega_i, \omega_{-i}), \omega_0, \omega_j\right) \middle| b, \omega_i\right]\Bigg\} \geq z_i. \end{aligned} \tag{C9}$$

We can maximize the expected revenue by letting z_i satisfy (C9) with equality for each $i \in N$. Accordingly, the central planner can receive from each agent i the expected value given by

$$\min\left\{E[v_i(g(\omega),\omega_0,\omega_i)|b]-c_i(b_i),\right.$$

$$\min_{\omega_i\in\Omega_i} E_{\omega_{-i}}\left[\sum_{j\in N\cup\{0\}} v_j\left(g(\omega_i,\omega_{-i}),\omega_0,\omega_j\right)\middle| b,\omega_i\right]$$

$$\left.-E\left[\sum_{j\in N\cup\{0\}\setminus\{i\}} v_j\left(g(\omega),\omega_0,\omega_j\right)\middle| b\right]\right\},$$

which implies (C5), where we add $E[v_0(g(\omega),\omega_0)|b]$.

Q.E.D.

Clearly, R^{EAIR} is equal to the maximized expected social welfare. From the relative strength of the incentive compatibility constraints, it is also clear that

$$R^{EPIR} \le R^{IIR,EAIR} \le R^{IIR},$$

and

$$R^{IIR,EAIR} \le R^{EAIR}.$$

Whether R^{IIR} or R^{EAIR} is larger depends on the situation.

The following proposition indicates that, with the constraints of EPIR, it is generally difficult for the central planner to achieve full efficiency without deficits:

Proposition 5: *Suppose that (b,g) is fully efficient, B is rich at b, and we have private values. Suppose also that $\sum_{i\in N} c_i(b_i)>0$ and there exists a null state $\underset{\sim}{\omega}=(\underset{\sim}{\omega}_0,\ldots,\underset{\sim}{\omega}_n)\in\Omega$ in the sense that $v_0(a,\underset{\sim}{\omega}_0)=0$ for all $a\in A$, and for every $i\in N$, $v_i(a,\omega_0,\underset{\sim}{\omega}_i)=0$ for all $a\in A$ and $\omega_0\in\Omega_0$. Then, with EPIR, the central planner has a deficit in expectation: $R^{EPIR}<0$.*

Proof: If $R^{EAIR}<0$, the conclusion would be immediate from the fact that $R^{EPIR}\le R^{EAIR}$. Suppose $R^{EAIR}\ge 0$. Because for $(\underset{\sim}{\omega}_0,\underset{\sim}{\omega}_1,\cdots,\underset{\sim}{\omega}_n)$, $\sum_{i\in N\cup\{0\}} v_i(a,\underset{\sim}{\omega}_0,\underset{\sim}{\omega}_i)=0$ holds for all $a\in A$, we have

$$\min_{\omega \in \Omega} \sum_{i \in N \cup \{0\}} v_i \left(g(\omega), \omega_0, \omega_i\right) \leq 0.$$

It follows from $R^{EAIR} \geq 0$ and $\sum_{i \in N} c_i \left(b_i\right) > 0$ that we have

$$E\left[\sum_{i \in N \cup \{0\}} v_i \left(g(\omega), \omega_0, \omega_i\right) \middle| b\right] \geq \sum_{i \in N} c_i \left(b_i\right) > 0.$$

From these observations,

$$R^{EPIR} = n \min_{\omega \in \Omega} \sum_{j \in N \cup \{0\}} v_j \left(g(\omega), \omega_j\right) - (n-1) E\left[\sum_{j \in N \cup \{0\}} v_j \left(g(\omega), \omega_0, \omega_j\right) \middle| b\right]$$
$$< 0.$$

Q.E.D.

If $R^{EAIR} < 0$, the conclusion is immediate from $R^{EPIR} \leq R^{EAIR}$. Suppose $R^{EAIR} \geq 0$. It follows from $\sum_{i \in N} c_i(b_i) > 0$ that the second term of R^{EPIR} in (C2) is negative. Due to the presence of the null state, the first term of R^{EPIR} in (C2) is non-positive. Accordingly, R^{EPIR} is negative.

By replacing EPIR with weaker participation constraints, such as IIR and EAIR, and adding some restrictions, it becomes much easier for the central planner to achieve full efficiency without deficits.

We say that we have *conditionally independent types* if for every $\tilde{b} \in B$, there exists a marginal type distribution of agent i, $f_i(\cdot\,|b_i)$ such that for all $\omega \in \Omega$, we have

$$f\left(\omega \middle| \tilde{b}\right) = \prod_{i \in N \cup \{0\}} f_i\left(\omega_i \middle| \tilde{b}\right),$$

When we have conditionally independent types, once we fix an action profile, ω_i provides no information for predicting ω_{-i}.

The following proposition states that, when we have conditionally independent types, the central planner can earn the same expected revenue as in the case of observable actions:

Proposition 6: *Suppose that (b, g) is fully efficient, B is rich at b, and we have private values and conditionally independent types. Then, the expected revenue achieved by any VCG mechanism that satisfies IIR and EAIR is less than or equal to $R^{IIR,EAIR}$. Furthermore, there exists a pure-VCG mechanism that satisfies IIR and EAIR and achieves $R^{IIR,EAIR}$.*

Proof: We have already shown that there is a pure-VCG mechanism that achieves $R^{IIR,EAIR}$. Assuming conditionally independent types and private values, we consider an arbitrary VCG mechanism (g, x) such that for each $i \in N$, there exists $y_i : \Omega_{-i} \to R$ such that

$$x_i(\omega) = -\sum_{j \in N \cup \{0\} \setminus \{i\}} v_j\big(g(\omega), \omega_0, \omega_j\big) + y_i(\omega_{-i}) \text{ for all } \omega \in \Omega.$$

EAIR implies

$$E[y_i(\omega_{-i})|b] \leq E\left[\sum_{j \in N \cup \{0\}} v_j\big(g(\omega), \omega_0, \omega_j\big)\middle| b\right] - c_i(b_i),$$

while IIR requires

$$E[y_i(\omega_{-i})|b, \omega_i] \leq E_{\omega_{-i}}\left[\sum_{j \in N \cup \{0\}} v_j(g(\omega_i, \omega_{-i}), \omega_0, \omega_j)\middle| b, \omega_i\right]$$

for all $\omega_i \in \Omega_i$, or, equivalently,

$$E[y_i(\omega_{-i})|b] \leq \min_{\omega_i \in \Omega_i} E_{\omega_{-i}}\left[\sum_{j \in N \cup \{0\}} v_j\big(g(\omega_i, \omega_{-i}), \omega_0, \omega_j\big)\middle| b, \omega_i\right].$$

Here, we used the fact that $E[y_i(\omega_{-i})|b] = E[y_i(\omega_{-i})|b, \omega_i]$ because of conditional independence. Accordingly, we have

$$
\begin{aligned}
&E[y_i(\omega_{-i})|b] \\
&\leq \min\left\{E\left[\sum_{j\in N\cup\{0\}} v_j\big(g(\omega),\omega_0,\omega_j\big)\middle| b\right] - c_i(b_i),\right. \\
&\qquad\qquad \left.\min_{\omega_i\in\Omega_i} E_{\omega_{-i}}\left[\sum_{j\in N\cup\{0\}} v_j\big(g(\omega_i,\omega_{-i}),\omega_0,\omega_j\big)\middle| b,\omega_i\right]\right\},
\end{aligned}
$$

which implies

$$
E\left[v_0(g(\omega),\omega_0) + \sum_{i\in N} x_i(\omega)\middle| b\right] \leq R^{IIR,EAIR}.
$$

Q.E.D.

One may expect that the central planner will receive a greater expected revenue than $R^{IIR,EAIR}$ once we assume that the actions are fixed or publicly observable. However, Proposition 6 indicates that, with conditionally independent types, no VCG mechanism is able to achieve the expected revenue greater than $R^{IIR,EAIR}$. This implies that $R^{IIR,EAIR}$ is the upper-bound of the expected revenue irrespective of the requirement of inducibility.[20]

Proposition 6 is related to the observation from a classical principal–agent model. When both the principal and the agent are risk-neutral, one of the revenue-maximizing contracts (which achieves the first-best of the principal) is to sell the company to that agent—that is, to give the entire outcome (externalities to the principal) in exchange for a fixed constant fee (Harris and Raviv, 1979). By regarding pure-VCG mechanisms as an extension of such selling-out contracts, Proposition 6 indicates that selling the company is the *unique* revenue-maximizing contract when the action space is rich at the efficient action profile and the agent has unlimited liability (which corresponds to EAIR and IIR).

[20] In the main body of this paper, we assume finite state spaces. Accordingly, some non-VCG mechanisms may be allocatively efficient and incentive compatible. However, if we take a fine grid of the states to make the state space closer to the continuum (with which the regularity condition of the Green–Laffont–Holmström theorem is satisfied), only the mechanisms that are close to VCG can be incentive compatible. In such a case, the maximum possible revenue is also close to $R^{IIR,EAIR}$.

It is widely accepted that efficiency is achievable through a VCG mechanism without running expected deficits if inducibility is not required. With conditionally independent types and some moderate restrictions, the non-negativity of $R^{IIR,EAIR}$ is guaranteed as follows:

Proposition 7: *Assume the hypotheses of Proposition 4, conditionally independent types, and the following conditions:*

Non-Negative Valuation: *For every* $i \in N \cup \{0\}$ *and* $\omega \in \Omega$*,*

$$v_i(g(\omega), \omega_0, \omega_i) \geq 0.$$

Non-Negative Expected Payoff: *For every* $i \in N$*,*

$$\text{(C10)} \qquad E[v_i(g(\omega), \omega_0, \omega_j)|b] - c_i(b_i) \geq 0.$$

With IIR and EAIR, the central planner has non-negative expected revenue: $R^{IIR,EAIR} \geq 0$.

Proof: From conditionally independent types, non-negative valuations, and null state, it follows that, for every $i \in N$ and $\omega_i \in \Omega_i$,

$$\begin{aligned}
& E_{\omega_{-i}}\left[\sum_{j \in N \cup \{0\}} v_j\left(g(\omega_i, \omega_{-i}), \omega_0, \omega_j\right) \middle| b\right] \\
&= E_{\omega_{-i}}\left[\max_{a \in N} \sum_{j \in N \cup \{0\}} v_j\left(a, \omega_0, \omega_j\right) \middle| b\right] \\
&\geq E_{\omega_{-i}}\left[\max_{a \in N} \sum_{j \in N \cup \{0\} \setminus \{i\}} v_j\left(a, \omega_0, \omega_j\right) \middle| b\right] \\
&= E_{\omega_{-i}}\left[v_i(g(\underline{\omega}_i, \omega_{-i}), \omega_0, \underline{\omega}_i) + \sum_{j \in N \cup \{0\} \setminus \{i\}} v_j\left(g(\underline{\omega}_i, \omega_{-i}), \omega_0, \omega_j\right) \middle| b\right],
\end{aligned}$$

which implies

$$\min_{\omega_i \in \Omega_i} E_{\omega_{-i}} \left[\sum_{j \in N \cup \{0\}} v_j \left(g(\omega_i, \omega_{-i}), \omega_0, \omega_j \right) \middle| b \right] = E_{\omega_{-i}} \left[\max_{a \in A} \sum_{j \in N \cup \{0\} \setminus \{i\}} v_j \left(a, \omega_0, \omega_j \right) \middle| b \right].$$

Therefore, we have

$$\min_{\omega_i \in \Omega_i} E_{\omega_{-i}} \left[\sum_{j \in N \cup \{0\}} v_j \left(g(\omega_i, \omega_{-i}), \omega_0, \omega_j \right) \middle| b \right] - E \left[\sum_{j \in N \cup \{0\} \setminus \{i\}} v_j \left(g(\omega), \omega_0, \omega_j \right) \middle| b \right] \geq 0$$

From the assumption of non-negative expected payoffs, we have

$$E[v_i(g(\omega), \omega_0, \omega_i) | b] - c_i(b_i) \geq 0$$

From these observations, for every $i \in N$

$$\min \Bigg\{ E[v_i(g(\omega), \omega_0, \omega_i) | b] - c_i(b_i),$$

$$\min_{\omega_i \in \Omega_i} E_{\omega_{-i}} \left[\sum_{j \in N \cup \{0\}} v_j \left(g(\omega_i, \omega_{-i}), \omega_0, \omega_j \right) \middle| b \right]$$

$$- E \left[\sum_{j \in N \cup \{0\} \setminus \{i\}} v_j \left(g(\omega), \omega_0, \omega_j \right) \middle| b \right] \Bigg\} \geq 0,$$

which, along with non-negative valuations, implies $R^{IIR,EAIR} \geq 0$.

Q.E.D.

Non-negative valuation excludes the case of bilateral bargaining addressed by Myerson and Satterthwaite (1983), where it is impossible for the central planner to achieve allocative efficiency without deficits. Non-negative expected payoff excludes the case of opportunism in the hold-up problem, where the sunk cost $c_i(b_i)$ is so large that it violates inequality (C10). By eliminating these cases and replacing EPIR with IIR and EAIR, we can derive the possibility result in liability implied by Proposition 7.

Acknowledgements

We thank Gabriel Carroll, Philippe Jehiel, Fuhito Kojima, Benny Moldovanu, Julia Salmi, Takuo Sugaya, and Yuichi Yamamoto, and seminar participants at Hitotsubashi University, University of Tokyo, Osaka University, Kyoto University, Stanford University, Seoul National University, and Waseda University. All errors are ours.

Funding sources

This study was supported by Grant-in-Aid for Scientific Research (KAKENHI 25285059, 26J10076, 16H02009) from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese government. This study was supported by the Funai Overseas Scholarship, the E.K. Potter Fellowship, the JSPS Research Fellowship (DC1), and the CARF Research Fund of the University of Tokyo.

Declaration of interests

We have no relevant material or financial interests that related to the research described in this paper.

References

Arozamena, L., Cantillon, E., 2004. Investment incentives in procurement auctions. *Rev. Econ. Stud*. 71(1), 1–18. https://doi.org/10.1111/0034-6527.00273.

Arrow, K., 1979. The property rights doctrine and demand revelation under incomplete information in *Economics and Human Welfare*, ed. by M. Boskin, New York: Academic Press, 24–39.

Athey, S., Segal. I., 2013. An efficient dynamic mechanism. *Econometrica*, 81(6), 2463–2485. https://doi.org/10.3982/ECTA6995.

Bergemann, D., Välimäki, J., 2002. Information acquisition and efficient mechanism design. *Econometrica*, 70(3), 1007–1033. https://doi.org/10.1111/1468-0262.00317.

Carroll, G. 2015. Robustness and linear contracts. *Am. Econ. Rev.* 105(2), 536–563. https://doi.org/10.1257/aer.20131159.

Clarke, E.H., 1971. Multipart pricing of public goods. *Public choice*, 11(1), 17–33. https://doi.org/10.1007/BF01726210.

Crémer, J. McLean, R.P. 1985. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2), 345–361.

Crémer, J., McLean, R.P., 1988. Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6), 1247–1257. https://doi.org/10.2307/1913096.

Dasgupta, P., Maskin, E. 2000. Efficient auctions *Q. J. Econ.* 115(2), 341–388. https://doi.org/10.1162/003355300554755.

D'Aspremont, C., Gerard-Varet, L.-A. 1979. Incentives and incomplete information *J. Public Econ*. 11(1), 25–45. https://doi.org/10.1016/0047-2727(79)90043-4.

Green, J. R., Laffont, J.-J. 1977. Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, 45(2), 427–438.

Green, J. R., Laffont, J.-J. 1979. Incentives in public decision making, Amsterdam: North-Holland.

Groves, T., 1973. Incentives in teams. *Econometrica*, 41(4), 617–631.

Harris, M., Raviv, A., 1979. Optimal incentive contracts with imperfect information. *J. Econ. Theory*, 20(2), 231–259.

Hatfield, J. W., Kojima, F., Kominers, S.D., 2015. Strategy-proofness, investment efficiency, and marginal returns: An equivalence. Working paper.

Hausch, D. B., Li, L., 1991. Private values auctions with endogenous information: Revenue equivalence and non-equivalence. Working paper.

Hausch, D. B., Li, L., 1993. A common value auction model with endogenous entry and information acquisition. *Economic Theory*, 3(2), 315–334. https://doi.org/10.1007/BF01212920.

Holmström, B. 1979. Groves' scheme on restricted domains. *Econometrica*, 47(5), 1137–1144. https://doi.org/10.2307/1911954.

Holmström, B., Milgrom, P., 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, 55(2), 303–328. https://doi.org/10.2307/1913238.

Jehiel, P., Moldovanu, B., 2001. Efficient design with interdependent valuations. *Econometrica*, 69(5), 1237–1259. https://doi.org/10.1111/1468-0262.00240.

Krähmer, D., Strausz, R. 2007. VCG mechanisms and efficient ex ante investments with externalities. *Econ. Lett*. 94(2), 192–196. https://doi.org/10.1016/j.econlet.2006.06.028.

Legros, P., Matsushima, H. 1991. Efficiency in partnerships. *J. Econ. Theory* 55(2), 296–322. https://doi.org/10.1016/0022-0531(91)90042-3.

Li, S. 2017. Obviously strategy-proof mechanisms. *Am. Econ. Rev*. 107(11), 3257–3287. https://doi.org/10.1257/aer.20160425.

Maskin, E.S., 1992. Auctions and privatization in *Privatization*, ed. by H. Siebert, Kiel: Institut fur Weltwirschaften der Universitat Kiel, 115–136.

Matsushima, H., 1989. Efficiency in repeated games with imperfect monitoring. *J. Econ. Theory*, 48(2), 428–442. https://doi.org/10.1016/0022-0531(89)90036-7.

Matthews, S. 1984. Information acquisition in discriminatory auctions in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, 49, 1477–1500.

Mezzetti, C. 2004 Mechanism design with interdependent valuations: Efficiency. *Econometrica*, 72(5), 1617–1626. https://doi.org/10.1111/j.1468-0262.2004.00546.x.

Myerson, R.B., Satterthwaite, M.A. 1983. Efficient mechanisms for bilateral trading. *J. Econ. Theory*, 29(2), 265–281. https://doi.org/10.1016/0022-0531(83)90048-0.

Noda, S. 2019. Full surplus extraction and within-period ex post implementation in dynamic environments. *Theor. Econ.* 14(1), 39–69. https://doi.org/10.3982/TE2226.

Obara, I., 2008. The full surplus extraction theorem with hidden actions. *B.E. J. Theor. Econ*. 8(1), 1–28. https://doi.org/10.2202/1935-1704.1399.

Persico, N. 2000. Information acquisition in auctions. *Econometrica*, 68(1), 135–148. https://doi.org/10.1111/1468-0262.00096.

Rogerson, W.P. 1992. Contractual solutions to the hold-up problem. *Rev. Econ. Stud.* 59(4), 777–93. https://doi.org/10.2307/2297997.

Stegeman, M. 1996. Participation costs and efficient auctions. *J. Econ. Theory* 71(1), 228–259. https://doi.org/10.1006/jeth.1996.0116.

Tan, G. 1992. Entry and R&D in procurement contracting. *J. Econ. Theory* 60(1), 41–60. https://doi.org/10.1016/0022-0531(92)90100-V.

Tomoeda, K. 2019. Efficient investments in the implementation problem. *J. Econ. Theory* 182, 247-278. https://doi.org/10.1016/j.jet.2019.04.009.

Vickrey, W. 1961. Counterspeculation, auctions, and competitive sealed tenders. *J Finance* 16(1), 8–37. https://doi.org/10.1111/j.1540-6261.1961.tb02789.x.

Williams, S.R., Radner, R. 1995. Efficiency in partnerships when the joint output is uncertain in *The Economics of Informational Decentralization*, ed. by J. O. Ledyard, Boston: Kluwer, 79–99.

Wilson, R., 1987. Game theoretic analysis of trading processes in *Adv. Econ. Theory* ed. by T. Bewley, Cambridge University Press.